Nordic genomic prediction for crossbred dairy females

***A. Fogh*[1]*, Vahlsten T.,*[2] *Markey C.*[3] *Byskov K.*[1]*, Liu H.* [1] *and Aamand G.P.*[4]**

[1] *SEGES Innovation, Udkærsvej 15, 8200 Aarhus N, Denmark*
[2] *Faba co-op, P.O. BOX 40, FIN-01301 Vantaa, Finland*
[3] *Växa, Box 288, 75105 Uppsala, Sweden*
[4] *Nordic Cattle Genetic Evaluation, 8200 Aarhus N, Denmark*

Abstract

Genomic predictions have been applied in dairy cattle for more than a decade with great success, but genomic estimated breeding values (GEBV) are not widely available for crossbred dairy females. In the Nordic countries NAV have implemented GEBV for female crosses between Holstein, RDC and Jersey. A breed of origin model (BOM) exploiting breed origin of the alleles (BOA) has been chosen, where SNP solutions from genomic prediction in the three pure breeds are utilized. Genotypes are imputed and phased using FImpute. Assignment to breed of origin is done by the AllOr program. GEBVs are calculated for the same main traits as purebred animals. Estimation of GEBVs for cross breeds include the following steps: 1) Express the SNP solutions from official pure breed genomic evaluations to phenotypic scale. 2) Calculate direct genomic value (DGV) by multiplying genotypes by purebred SNP solutions considering BOA. 3) Rescale the polygenic effect from genotyped purebred ancestors to phenotypic scale if available 4) Calculate polygenic effect for each animal. 5) Calculate GEBV from DGV and polygenic effect considering the genetic difference between breeds. 6) Combine breeding values for main trait groups in a total merit index (NTM). 7) Standardize GEBVs to an average of 0 for NTM and 100 for all other traits for animals in the base population. Animals in the defined genetic base population are genotyped crossbred females of 1 to 7 years of age at the publication date. Implementation of GEBVs for crossbred females in the Nordic countries give farmers the same opportunities to do within herd selection of their genomic tested crossbred heifers as of purebred heifers. Hence, this will be an important strategic tool at herd level to optimize the use of sexed semen and beef semen.

Key words: Crossbreeding, Genomic evaluation, dairy cattle

Introduction

Genomic predictions have been applied in dairy cattle for more than a decade with great success, but genomic estimated breeding values (GEBV) are not widely available for crossbred dairy cows. GEBVs for crossbred females will give farmers the same opportunities to do within herd selection of their genomic tested crossbred heifers as of purebred heifers. Hence, this will be an important strategic tool at herd level to optimize the use of sexed semen and beef semen.

Materials and Methods

Data:

In total 7.000 Nordic crosses between Holstein, RDC and Jersey were used to develop the genomic EBV’s.

Methods:

Genotypes are imputed and phased using FImpute (Sargolzaei et al., 2014). In total there were used 56K SNP. Assignment to breed of origin is done by the AllOr program (Eiríksson et al 2021).

GEBV's are calculated for 15 composite traits using the below model:

$$GEBV_{BOM,i} = \sum_{b=1}^{N_b} (v_b'(w_{i,1} \circ s_{1,i,b)} + v_b'(w_{i,2} \circ s_{2,i,b})) + \mu_b \frac{\sum s_{1,i,b} + \sum s_{2,i,b}}{2m} + a_i$$

Where:

- $w_{i,j}$ contains haplotype j coded as 0 and 1 for the alternative alleles
- v_b is a vector of marker effects for breed b
- $\circ$ is element vise multiplication
- $s_{j,i,b}$ is a vector of breed of origin indication for allele j of animal i to breed b, with 1 for alleles assigned to breed b and 0 for alleles assigned to other breeds and proportional values for alleles that could not be assigned
- μ_b is the intercept, accounting for difference in breed averages for breed b
- m is the number of markers
- a_i is a residual polygenic effect

The pedigree is traced back for 5 generations for the genotyped crossbred animals. The crossbred animals are only included if sire and maternal grandsire is either Holstein, Jersey or RDC. Given that, it follows that the dam is either a Holstein, Jersey, RDC or crossbred. Genotypes are extracted on crossbred cows and their genotyped purebred ancestors. Genotypes are imputed and phased using F-impute (Sargolzaei et al., 2014). Assignment to breed of origin is done by the Allor program (Eiríksson et al 2021). Animals where >10% of the alleles could not be assigned are excluded from the calculation of GEBVs.

SNP solutions for main trait groups, linear conformation traits and yield traits from the pure breed genomic prediction of RDC, Holstein and Jersey are used for all traits, except fertility and Saved feed where respectively interval from 1. to last insemination and maintenance is used. The SNP solutions from official genomic evaluation for Holstein, RDC and Jersey are on breed specific index scales. The SNP solutions are converted to original phenotypic scale by multiplying SNP solutions by the phenotypic values of +1 index unit to make them comparable across breeds. If a trait consists of more different subtraits, the subtrait with the highest correlation to the main trait across breeds was chosen.

Breed specific marker sets are used in the pure breed genomic predictions. In the cases where no SNP solution was available for a particular marker in a breed, the effect of the marker was set to zero.

Polygenic effect (ai) is calculated from genotyped purebred ancestors. If polygenic effect is available for a parent or parents, it is multiplied by 0.5 and included. If it is not available for one or more parents, then the grandparents are considered and the effect multiplied by 0.25, and great-great-grandparents in the same manner. Polygenic effects are also multiplied by the phenotypic values of +1 index unit to make them comparable across breeds.

The intercept terms, μ_b is taking the genetic difference between the breeds into account. Breed differences are calculated relative to HOL. Average of within country breed differences for RDC vs HOL and JER vs HOL (only Denmark) for birthyear class 2016 was used. For main trait consisting of more subtraits average of each subtrait was calculated and relative to the economic weight they were modified to breed difference for the subtrait with the highest correlation to the main trait. Breed solutions are corrected for average DGV converted to phenotypic scale from the purebred evaluations to account for different breed levels of DGV in birthyear class 2016.

GEBVs of all animals are adjusted in such a way that animals in the defined rolling genetic base population have an average index of 100.

For all traits, a rolling genetic base consisting of females from Denmark, Sweden and Finland that are 1-7 years of age at the date of publication is applied. Standard deviations of GEBVs for all animals are standardized with scaling factors used for HOL.

NTM is calculated by multiplying each standardized sub-index by a weighting factor. Weighting factors applied to HOL is used for the crossbreed evaluation.

The mean NTM in the rolling genetic base population is not adjusted, because it is based on adjusted EBVs for all sub-indices.

Results & Discussion

The Breed Origin Model is chosen due to validation results from the DairyCross project that shows that this model provides more accurate GEBV's than a Breed Proportion Model (Eiríksson et al 2022).

The developed GEBV's are validated by comparing them to pedigree EBV's. Genotyped crossbred animals are grouped within herd in two groups (HIGH vs LOW) based on either GEBV's or pedigree EBV. Only herds with at least 20 genotyped crossbred cows with phenotypic records for protein yield were included. Differences in protein yield between groups across herds (72 herds for 1. lact. and 68 herd for 2. lact.), are shown in figure 1.

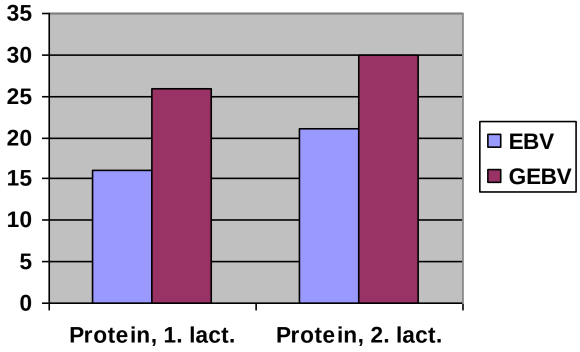

Figure 1. Difference between HIGH and LOW group within herd based on either pedigree EBV or GEBV.

Results show that GEBV is a better predictor of phenotypic performance than pedigree EBV.

Conclusions

GEBVs are a good management tool for farmers to assist decisions about which cows and heifers should be inseminated with either sexed semen or which should be inseminated with beef semen.

Acknowledgments

This work is based on results from the DairyCross project funded by Green Development and Demonstration Programme.

References

Eiríksson, J.H., Karaman, E., Su, G. et al. Breed of origin of alleles and genomic predictions for crossbred dairy cows. Genet Sel Evol 53, 84 (2021).

Eiríksson, J.H., Byskov, K., Su, G. et al. Genomic predictions for crossbred dairy cows by combining solutions from purebred evaluation based on breed origin of alleles. J. Dairy Sci., 105 (2022), pp. 5178-5191.

Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478.